

非定常多腕バンディットにおける社会的学習エージェントの分布則

北里大理^A 信州大理^B
 守 真太郎^A, 中山 一昭^B

Distribution of social learning agents in restless multi-armed bandit

^AKitasato Univ., ^BShinshu Univ
 S. Mori^A and K. Nakayama^B

他個体の選択や振る舞いを参考にして選択、行動を行う情報伝達、学習を社会的学習と呼び、自力で情報を探す個人的学習と比較して得られる情報の精度が悪くなくコストも低いことから広く使われている。本講演では非定常多腕バンディットにおいて、他のエージェントが持つレバー情報をコピーする社会的学習エージェントの統計物理的な解析結果について報告する ([1] preprint arXiv:1607.01944).

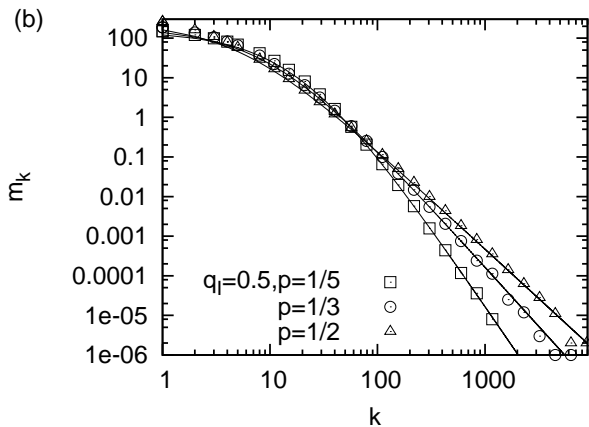


図1 エージェント数 k のレバーの分布関数 m_k . m_k は $N, M \rightarrow \infty$ の極限で Yule 分布に従う. $m_k \sim k^{-1-\gamma}$, $\gamma = 1 + \frac{(1-p)q_I}{pq_O}$.

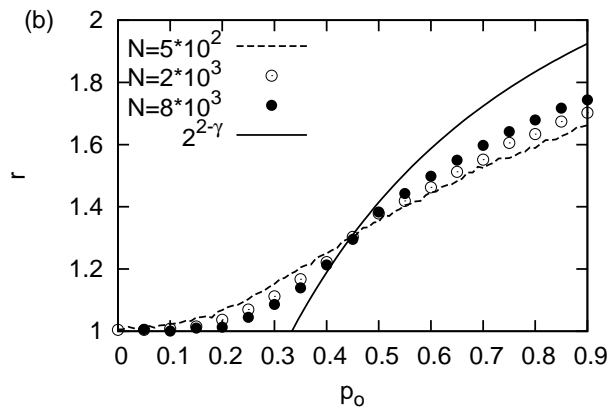


図2 N_1 のゆらぎ (個体あたりの分散) の $N, 2N$ で比 r の p 依存性をプロット.

エージェント数 N 、リターンのあるレバー数 M 、リターンのないレバー、1 ターンにランダムに選ばれたエージェントが以下の選択を行う。すべてのリターンのあるレバーが独立に確率 q_c/N でリターンのないレバーに変化し、リターンのあるレバーが追加され M は一定とする。エージェントは、リターンのあるレバーを知っていればレバーを引き、知らなければ確率 p で他のエージェントの知っているレバーを一本ランダムに選択して情報をコピー (成功確率 q_O)、確率 $1-p$ で自力でリターンのあるレバーを探索する (成功確率 q_I)。エージェントのレバーへの分布は $N, M \rightarrow \infty$ で Yule 分布に従う。また、リターンのあるレバーを知っているエージェント数 N_1 のゆらぎ (一個体あたりの分散) が $p_c = \frac{q_I}{q_I + q_O}$ を境に有限な相と N とともに発散する相の間の相転移を示す。